EPO-DG 1

2 3 -11- 2007

D1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁷: G06K 9/80

(21) International Application Number: PCT/US00/32885

(22) International Filing Date: 4 December 2000 (04.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/453,730   2 December 1999 (02.12.1999)   US

(71) Applicant *(for all designated States except US)*: KETEL, LLC [US/US]; 180 West End Avenue, 10D, New York, NY 10023 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: AMMAR, Maan [SY/SY]; P.O. Box 10650, Damascus (SY).

(74) Agent: HANDAL, Anthony, H.; Handal & Morofsky, 80 Washington Street, Norwalk, CT 06854 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *With amended claims and statement.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD AND APPARATUS FOR VERIFICATION OF SIGNATURES
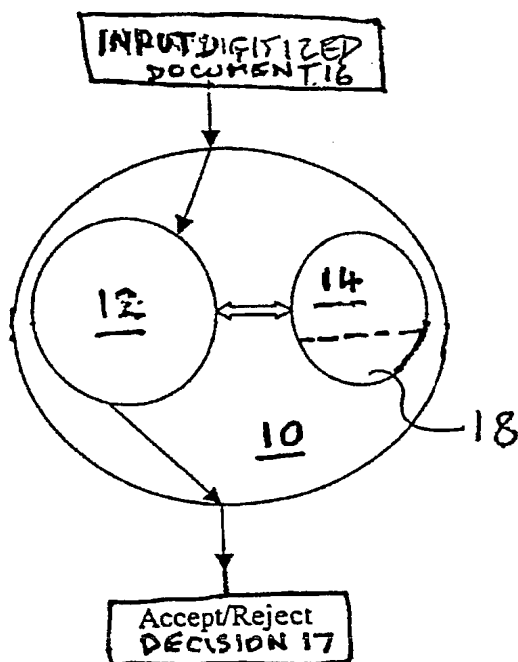
(57) Abstract: An automatic signature verification system (10) is set forth that utilizes a main routine (12) for comparing signatures using forensic hand writing methodology. A secondary program (14) is used to modify the algorithms used by the main program (12) for making adjustments thereto based on either additional data consisting of a plurality of genuine or authenticated signatures or changes in a persons signature due to aging or some other physical change resulting in a change in signature features. Over seventy thousand signatures can be verified in one hour using a single personal computer. In addition, authentic signature data having a resolution as low as 80dpi can be used.

WO 01/41051 A1

METHOD AND APPARATUS FOR VERIFICATION OF SIGNATURES

BACKGROUND OF THE INVENTION

1. Field Of The Invention

The present invention relates to a computerized system for automatically authenticating signatures. More

5   particularly, it relates to a system utilizing forensic methods of analysis while fully automating signature verification thereby reducing the time necessary for verification of a signature to a fraction of a second using ordinary personal computers.

2. Background Information

American businesses and households write over 60 billion checks per year. Banks and their customers lose

10 nearly ten billion dollars (Bank Administration Institute reports) to check fraud due to shortcomings during processing of those checks.

The Office of the Comptroller of the Currency (OCC) has indicated that check fraud is one of the largest challenges facing financial institutions. Modern technology allows more accurate forgeries to occur, making detection more difficult. In addition, technology allows more realistic counterfeit checks and false

15 identification that can be used in combination with forged signatures to defraud banks.
Many difficulties have been encountered in applying automatic signature verification systems to current data environments within banks. There is a need to verify signatures on checks that have a low resolution or poor quality image. The time it takes to verify each check with current automatic computer programs is far too long.

20 The OCC advises banks to review checks ensuring that the handwriting or print styles are consistent and that there are no signs of erasure or alteration. Banks should also compare the endorsement signatures on items presented and compare the appearance of the presenter with the signature and picture on their identification.

One commercial automatic signature verification system used by banks is provided by SOFTPRO under

25 the trademark SIVAL. This system employs a neural network training approach. The verification decision is performed using electronic comparison of images in which the set of parameters of the document's signature is compared with all parameter sets of the master signature of the account. Furthermore, every signature is assigned to one of the six defined agreement categories, AA to F, which can be subdivided into five categories if needed. The basis of both training and decision making (accepting the document

30 signature as genuine or rejecting it as a forgery) are drastically different from those of the present invention.

A bank or other financial institution may process millions of checks each day. Even with the help of computerized visual verification systems, only a small portion of the checks will have their signatures examined. This situation makes automatic verification a necessity to deal with large numbers of checks. The processing speed of a system becomes a critical factor for two reasons; (1) the time period allowed

5   for examining all of the checks is only two to four hours; and (2) the number of computers required to perform the work increases in proportion to the time t takes to verify each check.

The SOFTPRO system discussed previously has a speed of 1,800 documents per hour, using a Pentium II processor, which translates into 90,000 documents per hour using fifty computers. This number is far below the requirement of a large bank.

10   Another critical problem that has heretofore not been solved deals with the current database used by a bank. Some banks currently scan checks using hand scanners which have low resolution capabilities and poor quality signature card images. For example, documents scanned at 120 dpi and 80 dpi are available by the millions. Rescanning all of those documents is too time consuming and expensive. Current systems recommend a minimum of 200 dpi for scanned document data, and may be capable of using data

15   down to 150 dpi with a greatly reduced accuracy. The current systems with resolutions fail to provide meaningful results if data below 150 dpi are used.

A neural network trained using English style signatures can not be used to analyze Chinese style signatures. The network must be trained to handle type of writing, making it difficult to apply generalized software to different languages. Accordingly, there is a need for a system that can overcome the variations

20   in signature.

SUMMARY OF THE INVENTION

This invention solves a problem. It solves the problem of providing a signature verification system and method which can rapidly and reliably verifying a signature. Preferred embodiments solve further problems such as the problems of providing a signature verification system and method which can operate

25   at speeds sufficient for commercial banking purposes and with resolutions commonly available to banks. A further problem solved by preferred embodiments of the invention is that of verifying signatures written in a script other than the Latin alphabet, for example Chinese, Japanese or other pictographic scripts.

Accordingly, the invention provides, in one aspect, a system for automatically verifying signatures includes a program running on a personal computer using at least one authentic signature that has been scanned into

30   a genuine signature data base. A program caused the personal computer to run various algorithms to clean a digitized image of a target signature. The program then normalizes the image and makes Euclidian weighted measurements of forensic features of the target signature, such features being compared with those of the authentic signature features.

3

The algorithms used can be used to evaluate electronic image signatures having resolutions as low as eighty dpi or lower.

In contrast to known computerized systems that give the decision via comparison with only one genuine signature sample, the mechanism of giving the decision in the present invention is consistent with forensic
5   examination of handwriting samples. It is a well established fact in forensic examination that a person cannot write their signature the same way twice. Every sample is different from the other to some degree in that the values of the features of each signature vary as to relative size and two dimensional position, relative slants, curvature of letters, and the like. If the differences in the values of selected features of a target signature lie in a selected range of acceptance, as determined by the natural variations of the
10 signature of the specific person, the signature is accepted as genuine, otherwise, it is classified as an attempted forgery. In this context, the present invention uses image processing and pattern recognition techniques to implement forensic examination concepts during the decision making process by comparing the selected features of a target signature with a reference knowledge, or values, of the signature of the specific person obtained from a set of training or genuine signature samples. Training is done on two
15 levels: global, which includes the entirety of a genuine signature database of many persons, and personal, which includes a set of genuine signature samples of a specific person. While systems and methods of the present invention can use a set of genuine training signature samples for every person, it starts making decisions using only one reference signature sample using what is called here accelerated learning. In the beginning of the learning process, it uses the global knowledge as a starting point to give a decision using
20 only one reference signature. When new genuine signatures become available (from new checks or the like) the reference knowledge of a person's signature is updated until it becomes fully dependent on the genuine signatures of that person after only six genuine samples.

BRIEF DESCRIPTION OF THE DRAWINGS

Some embodiments of the invention, and of making and using the invention, as well as the best mode
25 contemplated of carrying out the invention, are described in detail below, by way of example, with reference to the accompanying drawings, in which like reference characters designate like elements throughout the several views, and in which:

Figure 1   is a functional block diagram of a signature verification system according to the invention illustrating the system's structure;
30 Figure 2   is an operational flow diagram of the automatic signature verification program shown in Figure 1;

Figure 3   shows thinned, binary and boundary-detected versions of a target signature image,

Figure 4   is a view showing a target signature image segmented in four quadrants using a gravity center method,

4

<div align="center">DESCRIPTION OF THE PREFERRED EMBODIMENTS</div>

Referring now to Fig. 1, the automatic signature verification system of the invention, referenced 10,

includes two main components, or subsystems, namely a signature processing program 12 and an automatic

performance evaluation program 14. A digitized document16, containing signature image data (or other

image data of interest) from a check, or the like, regarding a questioned signature, the authenticity of which

15   is to be verified by system 10, is presented for evaluation to signature processing program 12. Signature

processing program 12 looks to automatic evaluation program 14, which includes an authenticated

signature database 18, to provide information for a forensic evaluation of the questioned signature.

According to the outcome of the evaluation, system 10 either accepts the questioned signature as genuine,

20   or rejects it as a forgery, in decision step 17.

Image data regarding signatures accepted as genuine may be fed back into automatic evaluation program

14, if desired, and included in authenticated signature database 18, to enhance signature database 18 and

improve system performance, as described in more detail hereinbelow.

Referring now to Fig. 2, digitized document16 is supplied to signature verification system 10 for

25   verification of the document's signature. Digitized document 16 may be received directly from a scanner,

or the like, from storage in a local or remote database, or in essentially real time from a network, for

example the Internet. Digitized document 16 may comprise an image of a bank card signature record, a

check, or other financial instrument, or other document, or the like, bearing the questioned signature.

It will be understood that although check signature verification comprises a preferred embodiment of the

30   invention, the invention can with advantage be applied to verification of signatures on other documents or

media, for example, credit card sales slips, government forms and the like, or to the verification of other

desired indicia, as may be understood from the teaching herein.

Initially, signature processing program 12 treats digitized document16 to a signature area extraction 20

which extracts from document 16 signature image data specified in a signature area specified by coordinates in the particular document type. If the original signature image data is in gray scale form, the image from signature area 20 is manipulated, e.g. by thresholding, to obtain a binary signature image 22 of the questioned signature.

5   Preferably, a signatory identifier, for example an account number on a check, is read by system 10 from digitized document 16 or elsewhere. Alternatively, the identifier can be manually posted or electronically transmitted to system 10. The identifier can be system-associated with signature image 22 and its processing and can be employed to retrieve personalized data regarding or derived from genuine signatures of the signatory to digitized document 16.

10   The preferred verification process of the invention continues with a preprocessing procedure 19 followed by a dissimilarity measurement procedure 21. Preprocessing procedure 19 prepares one or more versions of signature image 22 for dissimilarity determination in dissimilarity measurement procedure 21 by enhancing, cleaning, filtering orienting, positioning, or other such steps to present each signature image 22 in a standardized manner suitable for analysis of its characteristic features. Preferably, multiple versions
15   of signature image 22 are prepared, as will be described hereinbelow, providing different logical views of the questioned signature, for example, unmodified, thinned and boundary-detected images. Other possible views will be known or apparent to those skilled in the art. With advantage, from two to about ten, preferably no more than about six such image versions are prepared, while the three such image versions mentioned are preferred.

20   Dissimilarity measurement procedure 21 examines each image version for characteristic features and compares each individual characteristic feature with data regarding the feature obtained from database 18. The feature data is obtained using at least one, and preferably a number of genuine signatures, for example, from five to ten..

     After extraction from digitized document 16, and pursuant to preprocessing procedure 19, signature image
25   22 is subjected to an automatic cleaning step 24. Cleaning step 24 can, for example, be effected using connected-component labeling based cleaning, with horizontal and vertical line elimination taking place or in other ways. Thus for example, image components connected with other components are labeled and recognized as signature image components and are retained, while simple unconnected image elements may be discarded unless they meet criteria indicating they are components of the signature image. After
30   automatic cleaning step 24, the resulting data represents a clean binary signature image 26 with printed background characters, lines and noise eliminated. Cleaned binary signature image 26 is then prepared for feature extraction using forensic handwriting analysis techniques.

One version of cleaned signature image 26, an unmodified version 31, is passed directly to dissimilarity measurement procedure 21. A second version of cleaned signature image 26 is processed in a normalization procedure 33 to adjust the image to a system standardized configuration with regard to orientation, position, line width and so on. Normalization procedure 33 commences with an optional

5 orientation normalization routine 28 in which signature image 26 is normalized to a horizontal position. Orientation normalization routine 28 is used to reexamine a questioned signature 16 which has been found by system 10 to be a forgery, since a forgery finding may be caused by a significant change in the orientation of a genuine signature from a normal position.

The second versions of all cleaned signature images 26 are then subjected to a position normalization

10 routine 29 wherein signature image 26 is set to an origin of coordinates at a center of gravity of image 26. so that the authenticity decision is independent of the relative position of the signature image within the document signature area.

After position normalization 29, a third version of signature image 26 is sent for processing to a line thinning routine 30, proceeding on a parallel path to the second versions, before all three versions of

15 signature image 26 arrive at dissimilarity measurement procedure 21. Line thinning routine 30 provides a thinned signature image 32

Along a parallel path, second version, cleaned, position-normalized signature image 26, proceeds to size normalization routine 34 where the image is normalized with respect to writing line width and writing line size resulting in a size and line width normalized, or boundary detected image 40 Optionally, Debris or

20 artifacts are filtered out in an artifacts removal routine 36 in preparation for sending the signature image to the next step in its processing.

If desired, all three signature image versions 26, 32 and 40 can be size and position normalized for more accurate statistical analysis of their characteristic features against a reference database, in dissimilarity measurement procedure 21.

25 The three image signature image versions binary or unmodified image 26, thinned image 32 and boundary detected image 40, as illustrated in Figure 3, are next processed in dissimilarity measurement procedure 21 which begins with a segmentation and feature extraction routine 38.

During signature segmentation and feature extraction routine 38, for each of the three signature image versions 26, 32 and 40, the signature are is logically subdivided into a number of contiguous geometric

30 zones and the features in each zone are extracted. The number and shapes of the zones may be varied, provided that sufficient feature information can be extracted and processed quickly enough for the

purposes of the invention. For example, from 2-20, preferably from 2- 6 zones in the shape of rectangles, quadrants, sectors, bands, rings, combinations of these shapes, or more complex, interfitting shapes, can be employed. Also, multiple zone patterns may be employed sequentially, if desired, to provide different extracted feature views or mixes.

5   As shown in Figure 4, in one preferred embodiment, the signature image area of each version 26, 32 and 40 (26 is referenced by way of example) is segmented into four quadrants 46 about a center of gravity 47 of the signature image for one mode of feature extraction. As shown in Figure 5, the signature area is segmented into two vertical zones 48 using a baseline and into two horizontal zones 50 using a geometrical mid point 51.

10 The horizontal, positive, vertical, and negative slant features in each of quadrants 46 and zones 48 and 50 are computed and output as features F1, F2, ...Fn, referring to Figure 2. Extracted features 52 of a given signature image pixel, $p(i,j)$ as follows: for horizontally slanted $p(i,j+1)$; for positively slanted $p(i-1,j+1)$; for vertically slanted $p(i-1,j)$; and for negatively slanted $p(i-1,j-1)$ represented in Fig. 2. Extracted features 52 are analyzed on a global basis for the entire signature image and locally in each quadrant, each vertical
15 zone, and each horizontal zone thereby providing 120 extracted features 52, in the preferred example described herein, namely twelve global and one hundred and eight local features, each extracted feature 52 being measured on three image versions 26, 32, and 40. Feature extraction may be effected in accordance with the teaching of Ammar *et al.* in *Computer Recognition and Human Production of Handwriting* Eds. R. Plamondon, C.Y. Suen & M. L. Simner, World Scientific Publ. Co., 1989, pp. 61-76, ("Ammar *et al.*
20 1989" hereinafter), the disclosure of which is hereby incorporated herein by reference thereto. .

Two additional complex element features F121 and F122, are extracted by matching and mismatching, a horizontal profile of signature 26 with a mean horizontal profile of a reference or training sample comprising at least one genuine signature as described hereinbelow. Thus, in the preferred embodiment, one hundred and twenty simple features and two complex features 52 are extracted, F1-F122. Each feature
25 can then be individually weighted, pursuant to the invention, in feature weighting step 54 with weights W1, W2...Wn, as described hereinbelow. Preferably, the weightings are automatically selected by automatic evaluation program 14, as described in further detail below.

Weighted features F1-F122 are processed to provide a weighted Euclidean distance measure 56. A dissimilarity measure 60 is computed from weighted Euclidean distance measure 56. Optionally, but
30 preferably, the computation is effected using a desired feature set 58 selected automatically using automatic evaluation program 14 for dissimilarity measurements. Alternatively, all features F1-F122 may be employed.

Decision routine 62 then evaluates dissimilarity measure 60 against an adaptive threshold 63, which can be varied to enhance the quality of decision making, to verify signature image 22 as genuine or to reject it as a forgery.

Decision routine 62 determines an upper limit of the natural range of the dissimilarity measure for a
5   specific person, for example the signatory to digital document 16. The upper limit of the adaptive threshold is computed using parameters obtained from automatic evaluation program 14.

Dissimilarity measurements 60 give the distance relationship of selected features of questioned signature 16. This measure gives an indication of how far questioned signature 16 is from a set of training, or authentic, samples of the same person's signature. If all the training samples and questioned signatures are
10   copies of each other, the dissimilarity measure is zero. To the degree that the training samples differ from each other and the questioned signature differs from them, the dissimilarity measure rises proportionally to that degree. For a specific person's signature, there is a natural degree of variation in the values of the features of the samples of their signature such that there is a natural range of dissimilarity measurement of the person. If the dissimilarity measure of the questioned signature exceeds this natural range, the
15   questioned signature is judged to be an attempted forgery, otherwise, it is accepted as genuine. Computing this natural range is done with the help of automatic evaluation program 14, as described below.

As shown in Figure 6, signature verification program 12 can comprise the usual File, Window and Help commands, File having options for scanner selection and image acquisition. The "SamChecks" command provides access to a database of sample checks while customized maenus for database management and
20   cleaning are included.

Referring now to Fig. 7, the command menu of automatic evaluation program 14 is shown. Automatic evaluation program 14 is a powerful development tool for augmenting automatic signature verification algorithms. It enables evaluation of the performance of signature processing program 12 with respect to different image types, features, and decision-making methods. It also enables evaluation of the
25   performance with respect to some conditions, stability degree, for example.

When setting up or using the system in a manual mode the aforementioned evaluations can be done in one click for the whole signature database. In this way, it is easy to reach the best possible features, image types, and resolutions under any conditions.

In order to compute the percentage of correct acceptance and correct rejection of genuine and forgery
30   samples, respectively, or the two type error rates, the signature database used for evaluation must contain genuine and forgery samples of every person in the database. This is provided by collecting a signature

database of over one thousand signatures comprised of persons having different language backgrounds, such as Arabic, English, Korean and Japanese. The number of genuine samples and forgeries should be almost equal. With a database having these criteria, the minimum requirements for computing the parameters to be used in distance measure and adaptive decision threshold computations in automatic

5   signature verification system 10 are available.

Automatic evaluation program 14 can be run in the following three phases:

(1) Making Reference: For every person in the database, program 14 processes their genuine signatures, extracts features, computes reference statistics (mean and standard deviation of every feature of the one hundred and twenty two extracted features), and stores them in a corresponding file.

10   (2) Distance Measure: program 14 computes the distance measure (dissimilarity measure) for every genuine and forged signature in the signature database using weighted Euclidean distance routine 56 where the weights computed in (1) are the standard deviation of each feature. The distance measure for genuine samples and forgeries is computed as follows:

(a) Genuine signatures: For genuine signatures, use leave one out method, which is appropriate

15   even with a small number of reference samples. In this method, each reference sample is left out from the reference samples when its distance measure is computed. The reference statistics are computed from the remaining reference samples, and the distance measure of the left out sample is computed using these statistics for performance evaluation.

(B) Forged signatures: For the forged signatures, the reference statistics are those computed in (1).

20   The distance measures for forgery samples are computed using those statistics. Now, after computing the distance measure for every sample in the database, the performance evaluation is done as follows:

(3) Performance Evaluation: the decision making process is necessary for evaluation. The decision is made as follows: If the distance measure is larger than the verification threshold, the questioned signature is accepted as genuine, otherwise, it is rejected and classified as an attempted forgery.

25   Correct decisions: (1) If the questioned signature is genuine and accepted, it is a correct decision.

(2) If the questioned signature is a forgery, and rejected, it is a correct decision.

In order to explain the performance evaluation, first, the following definitions are introduced:

PCA: Percentage of Correct Acceptance. (Genuine accepted as genuine)

30   PCR: Percentage of Correct Rejection. (Forgery rejected and classified as a forgery)

SR: System Reliability. (PCA+PCR)/2 (average percentage of correct decisions).

In order to evaluate performance of the ASVR using a specified feature set, The present invention uses a threshold based decision for finding how the performance (percentage of correct decisions) for the whole database varies with decision threshold change. In other words, the upper limit of the natural variation

35   range of the dissimilarity measure of every person is marked by a threshold value (the verification

threshold). This threshold is computed using a formula containing a constant k. This constant is given different values from 0.1 to 3.0 in 0.1 increments producing thirty different verification thresholds. For every value of the thirty values of the verification threshold, PCA PCR and SR are computed on the whole signature data Using the resultant values, the PCA, PCR and SR curves are computed.

5   The SR curve always has a relative maximum (a peak) showing the maximum level of accuracy obtained using the signature data used and a specific feature set. Naturally, if the data changes, the result will change somewhat respectively, either positively or negatively. The peak may or may not be the same crossover (tradeoff) point of the three curves.

        If we choose the verification threshold to be the crossover point, then we choose the equal probability of

10   error for PCA and PCR. If we take into account the error rate, then:

        Type I error = 100 - PCA

        Type II error = 100 - PCR

        Average error rate = (Type I + type II)/2

15   Now, if we choose the verification threshold to correspond to a higher value than that which corresponds to the tradeoff point (equal two type errors point, or equal PCA and PCR point), then we will have a higher PCA and smaller PCR, due to more variations allowed in the input signature. The outcome of this higher threshold will not only be the acceptance of more genuine signatures for the same person, but the acceptance of more forgeries if they fall in the same range, too.

20   On the other hand, if we use lower threshold value, we will allow fewer variations of the signature; thus, a smaller number of genuine signatures will be accepted (less PCA), and more forgeries will be rejected (higher PCR). The threshold can be adjusted by the user to tighten or loosen the control in the general case. The individual accounts can be controlled by adjusting the learned parameter of the personalized thresholds.

25   The number of accepted forgeries depends on both the degree of stability of one's signature and the easiness of copying the signature. Such complexity is taken into account by computing decision parameters through a learning process of the entire database (global learning), which combines different types of signature shapes and degrees of stability as the appropriate ones. For some persons, however, this global threshold may not produce the best results.

30   The global threshold is not always the best threshold for everyone because it obviously is not customized to suit each individual. Such customization, however, may be closely attained by using the threshold which

corresponds to the crossover point. The present invention uses "personalized thresholds," which are computed using a formula containing the average and the standard deviation of the distances of the genuine samples of the specific person (local learning), in order to acquire his or her individual statistics. By modifying the global threshold to include the individual statistics, the software produces the best results.

5   Evaluation program 14 gives automatically the best global verification threshold (GVTH) which corresponds to the maximum value of SR curve.

The personalized verification threshold is obtained by dividing the GVTH by the mean distance of the genuine samples (training) of the same person. This method makes the verification threshold adaptive to the person's signature which can be experimentally determined to give the best result judged by program

10   14.

As described above, in order to be able to compute the distance measure it is desirable to have some training samples. This makes verifying with single reference signature impossible. However, this is the actual case in banks where one or two reference samples on the bank card are available. In order to overcome this dead lock with this method of distance measure and verification, the AL is developed for the

15   present invention.

In order to have meaningful learning about the person's signature, there five genuine signatures are desirale or more. In the actual case, this means that the ASV cannot make decision until four more genuine signatures collected from the checks of that person. Practically, there is no guarantee that all checks bear genuine signatures so that the ASVR can not work with single reference signature. Accelerated learning

20   developed in this invention solves this problem.

Accelerated learning is given this name because it accelerates the learning process before having five or more genuine signatures. It works as follows:
the signature database containing genuine and forgery samples of every person, is used with program 14 to get four classes of standard deviation of every feature used for verification. The four classes are

25   (Horizontally, Positively, Vertically and Negatively slanted signatures). Program 14 is run for every class of these signatures available in the signature database, and at the end of the run, the mean of the standard deviation of each feature for each class is recorded as the "learned standard deviation" of the feature so that the Euclidean distance measure equation can be executed using the learned standard deviation as a "weight" and the feature of the single reference signature as a "mean". Experiments can show that using

30   this accelerated learning method gives a surprisingly high percentage of correct decisions even in the case of single reference signature which may even be close to that obtained with the standard deviation values obtained from five or more training samples.

Referring to Fig. 8, the actual learning and decision making in the present invention include: (1) The single reference signature 70 is used to compute the feature values, step 72 and these values are used as the "mean" of the feature in the Euclidean distance measure equation 74.

(2) The standard deviation 76 learned by accelerated learning is used as the standard deviation (weight) of the feature in the Euclidean distance measure equation 74.

The distance measure used for verification is computed, step 78, using these values until five genuine signature obtained from the checks verified for that person.

Starting from the sixth genuine signature, the learned values of the mean and standard deviation of each feature is computed from the values of the features computed on the set of six training samples including the single reference signature itself. From this point on, the accelerated learning values are abandoned (not used for this person). This process is explained in Figure 8.

As new genuine signatures come from checks, the number of learning samples increases and the accuracy improves until ten genuine training samples are collected.

The feature set used by the present invention is selected automatically by program 14 as follows:

(1) the one hundred and twenty features are arranged in a one dimensional feature table so that the number of the table entry contains the features from one to its number, i.e., entry number six contains features (1,2,3,4,5,6), entry number four contains features (1,2,3,4), and so on.

(2) A circulant feature matrix is created from this one dimensional feature matrix so that the circulant matrix is nXn if the dimensions of the one-dimensional matrix is n.

(3) Program 14 is run for all entries of the two-dimensional feature matrix and for every entry selects the maximum SR.

(4) Finally, program 14 selects the entry which gives the maximum SR among all entries to determine the feature set used by the present invention (the best feature set).

Signature verification system 10 then uses the two complex features f121 and f122 to give the first decision concerning the questioned signature. If it is accepted, it is passed to the second stage which uses the best feature set selected by program 14.

The present invention can process an account with more than one signatory. First, the signature area is processed using vertical projections and connected-components labeling, with suitable rules to identify the presence of one, two or more signatures. If multiple signatures are recognized as being present, the signature area is segmented into individual questioned signature areas. Each questioned signature, in turn, is compared with the reference signatures of the authorized account signatories, using a distance measure

basis to identify a corresponding reference signature. Once the corresponding reference signature is identified, the questioned signature is verified by signature verification program 12 and accepted or rejected. The process is then repeated for the other questioned signatures.

5   Since the inventive signature verification system 10 evaluates signatures against a database of signature features, not by matching images and sub-images with each other, the verification time is very short compared with prior art image matching methods.

The systems and methods of the invention are intended for implementation on a computer having access to appropriate I/O devices, e.g. monitor, scanner, printer, communications devices and so on, as needed. While a personal computer is preferred at the present time, it will be understood that other computer or

10   intelligent information processing apparatus may be employed, including workstations, minicomputers, mainframes, distributed processing systems, a dedicated intelligent appliance, and so on. In typical implementations, a personal computer station running the programs employed in the invention will have access to a network, where desired or necessary resources are available to the personal computer, for example, scanners or scanned image libraries, signature and signature feature databases and so on.

15   Program software can be written to work under Microsoft WINDOWS (trademark) 95, 98, 2000, CE, NT or the like operating systems, and stored on physical media such as magnetic or optical disks, as is known in the art.. Preferably, the software programs of the invention are provided as DLL components that can be integrated into other software platform.

It will be understood that a computer station running signature processing program 12, with or without

20   performance evaluation program 14, comprises a novel signature verification machine capable of electronically evaluating questioned signature images and tagging or otherwise identifying each signature as being either genuine or a forgery, with good accuracy.

The systems and methods of the invention can, in preferred embodiments, operate effectively on a heterogenous mix of documents of differing quality, scanned at different resolutions, and signed in

25   different scripts, such as is the nature of the checks received for clearance by many banks around the world.

One example of the operation of evaluation program is illustrated in Figure 9. As shown in the example of Figure 8, PCA, the percentage of correct acceptances of questioned signatures increases sharply with increasing K from an unacceptably low value of about 22% when K equals 1, until the value

30   asymptotically approaches 100% when K equals 3. However, when K equals 3, the percentage of rejections, PCR, which is virtually 100% when K equals 1, becomes unacceptably high at 50. The system reliability curve indicates a trade-off between the two and peaks around an optimum value of K = about

1.8, where both the correct acceptances and correct rejections are at or above 80%. The balance between correct acceptances and rejections can be varied by varying K either side of the optimum value, as desired by the user. Thus, if the user prefers a high rate of correct acceptances, they may choose a value of K of about 2.3 yielding a 95% PCA and tolerating a PCR of 65%. Other choces will be apparent to those skilled

5   in the art.

Figure 10 illustrates an example of distance determinations obtainable with the invention in a test run on using check signatures provided by individuals identified on the X-axis. The distances obtained are shown on the Y-axis and it may be seen that in most cases the Forged signatures F, with data points "x" are more distant than the genuine signatures G with data points "•".

10   While illustrative embodiments of the invention have been described above, it is, of course, understood that various modifications will be apparent to those of ordinary skill in the art. Many such modifications are contemplated as being within the spirit and scope of the invention.

Claims: 15

1. A system for automatic signature authentication comprising a computer-implementable signature verification program to examine a digitized signature image and authenticate the signature as genuine or reject the signature as a forgery **characterized by** further comprising;

an automatic evaluation program to evaluate the results produced by the signature verification program and provide input to the signature verification program to enhance the results and an authentic signature data base for use by said automatic evaluation program.

2. A system according to claim 1 wherein the digitized signature image is scanned from a paper document **characterized in that** multiple versions of the signature image are generated and examined the results of which examination are statistically compiled to enhance the authentication decision.

3. A system according to claim 3 **characterized by** defining the signature image into segments, extracting signature features from individual segments and evaluating the extracted signature features against training data received from the evaluation program.

4. A system according to claim 3 **characterized in that** the extracted signature features are evaluated for their distance from a reference obtained from the evaluation program, the distance being computed from at least one genuine signature and being used to effect the authentication decision according to a threshold.

5. A system according to claim 4 **characterized in that** the threshold is adapted by the evaluation program to enhance the decision.

6. A system according to claim 4 **characterized in that** the threshold is adapted by the evaluation program according to new data from genuine signatures received by the evaluation program.

7. A system according to claim 3 **characterized in that** the evaluation program determines a feature set of said extracted signature features to be used by the signature verification program for the decision.

8. A system according to claim 1 **characterized by** cleaning the scanned signature image to remove printed background.

9. A system according to claim 1 **characterized in that** a scanned image is manipulated by the signature verification program into at least one digitized image having recognizable features used by said verification program to compare to similar features of an select image in said authentic signature data base for determining if said scanned image is a genuine signature.

10. A system according to claim 1, 3, 5, 6, 7 or 8 characterized by comprising a bank check verification system.

11. A method of verifying hand written signatures comprising the steps of:

scanning a questioned signature to be examined for authenticity forming a questioned binary image, manipulating said binary image to produce a questioned signature feature set, comparing said questioned signature feature set to a feature set of a known authentic feature set to determine whether the questioned signature is genuine or a forgery, using an automatic learning program to modify algorithms used by the processing program to increase accuracy thereof.

17
AMENDED CLAIMS
[received by the International Bureau on 23 April 2001 (23.04.01);
original claims 1-11 replaced by new claims 1-39 (6 pages)]

1. A system for automatic signature authentication comprising:

 a) a computer-implementable signature processing program (12) to examine a digitized signature image (22) obtained from a document, optionally a check, signature image (22) purportedly being the signature of a specific person, to generate an authentication decision (17) authenticating the signature as genuine or rejecting the signature as a forgery; and

 b) an authenticated signature database (18) of signature features (F1..Fn) for use by signature processing program (12) in generating authentication decision (17);

 characterized by further comprising:

 c) an evaluation program (14) to evaluate the results produced by the signature verification program (12) and to provide update information to the signature database (18) for use by signature verification program (12) regarding examined signatures accepted as genuine by signature verification program (12).

2. A system according to claim 1 characterized in that the system comprises a global signature feature database including signature features of genuine and forged samples of signatures of persons other than the specific person's signature together with available genuine samples of the specific person's signature.

3. A system according to claim 2 characterized in that the global signature database is employed for signature recognition until the database includes a desired number of genuine samples of the specific person's signature.

4. A system according to claim 1 wherein the digitized signature image (22) is scanned from a paper document characterized in that multiple versions of the signature image (22) are generated and examined the results of which examination are statistically compiled to enhance the authentication decision (17).

5. A system according to claim 4 characterized in that digitized signature image (22) is cleaned prior to generation of multiple versions by removing printed background characters, noise and unconnected lines not meeting signature image component criteria.

6. A system according to claim 4 characterized in that the multiple versions of the signature image (26) comprise an unmodified version (31), a thinned version (32) and a boundary-detected version (40).

7. A system according to claim 4 characterized in that the multiple versions of the signature image (26) comprise an unmodified version (31), a thinned version (32) and a boundary-detected version (40) wherein the unmodified version (31) is obtained by cleaning signature image (26) and is passed directly to a dissimilarity measurement procedure (21) and wherein a second version of cleaned signature image (26) is

processed in a normalization procedure (33) to adjust the image to a system standardized configuration with regard to orientation and position, thinned version (32) and boundary-detected version (40) being generated from the normalized second version image.

8. A system according to claim 7 characterized in that thinned version (32) is generated from the normalized second version image in a line thinning routine (30) and boundary-detected version (40) is generated from the normalized second version image where the image is normalized with respect to writing line width and writing line size in size normalization routine (34) to provide boundary detected image (40).

9. A system according to claim 4, 5, 6, 7, or 8 characterized in that the signature image (22) is defined into segments, signature features (52) are extracted from individual segments and extracted signature features (52) are evaluated against training data received from the evaluation program (14).

10. A system according to claim 9 characterized in that the segments comprise a first set of four quadrants (46) defined from a center of gravity (47) of the signature image, a second independent set of two vertical zones on either side of the center of gravity (47) and a third independent set of two horizontal zones on either side of the center of gravity (47).

11. A system according to claim 10 characterized in that horizontal, vertical, positive and negative slant features in each said segment are computed and output as signature features ((F1..Fn) to the signature database (18).

12. A system according to claim 9, 10 or 11 characterized in that the extracted signature features are each analyzed globally for the complete signature and locally for each segment.

13. A system according to claim 9 or 10 characterized in that two complex element signature features (F121, F122) are extracted by matching and mismatching, a horizontal profile of signature image (26) with a mean horizontal profile of a training sample comprising at least one genuine signature.

14. A system according to claim 9 characterized in that the extracted signature features (52) are evaluated for their distance from a reference obtained from the evaluation program, the distance being computed from at least one genuine signature and being used to effect the authentication decision (17) according to a threshold (63).

15. A system according to claim 14 characterized in that the threshold (63) is adapted by the evaluation program (14) to enhance the decision.

AMENDED SHEET (ARTICLE 19)

16.A system according to claim 14 or 15 characterized in that the threshold (63) is adapted by the evaluation program (14) according to new data from genuine signatures received by the evaluation program (14).

17. A system according to claim 7 characterized in that the evaluation program (14) determines a feature
5   set of said extracted signature features (52) to be used by the signature verification program (12) for the decision.

18. A system according to claim 1 characterized by cleaning the scanned signature image (22) to remove printed background.

19.A system according to claim 1, 3, 4, 6 or 7 characterized by comprising a bank check signature
10   verification system.

20. A system for automatic signature authentication comprising:
     a)   a computer-implementable signature processing program (12) to examine a digitized signature image (22) obtained from a document, optionally a check, signature image (22) purportedly being the signature of a specific person, to generate an authentication decision (17) authenticating the
15   signature as genuine or rejecting the signature as a forgery; and
     b)   an authenticated signature database (18) of signature features (F1..Fn) for use by signature processing program (12) in generating authentication decision (17);
     characterized in that multiple versions of the signature image (22) are generated and examined, the results of which examination are statistically compiled to enhance the authentication decision (17).

20   21. A system according to claim 20 characterized in that the multiple versions of the signature image (26) comprise an unmodified version (31), a thinned version (32) and a boundary-detected version (40) wherein the unmodified version (31) is obtained by cleaning signature image (26) and is passed directly to a dissimilarity measurement procedure (21) and wherein a second version of cleaned signature image (26) is processed in a normalization procedure (33) to adjust the image to a system standardized configuration
25   with regard to orientation and position, thinned version (32) and boundary-detected version (40) being generated from the normalized second version image.

22. A system according to claim 21 characterized in that thinned version (32) is generated from the normalized second version image in a line thinning routine (30) and boundary-detected version (40) is generated from the normalized second version image where the image is normalized with respect to writing
30   line width and writing line size in size normalization routine (34) to provide boundary detected image (40) and characterized in that the signature image (22) is defined into segments, signature features (52) are

AMENDED SHEET (ARTICLE 19)

extracted from individual segments and extracted signature features (52) are evaluated against training data received from the evaluation program (14).

23. A system according to claim 22 characterized in that the segments comprise a first set of four quadrants (46) defined from a center of gravity (47) of the signature image, a second independent set of
5   two vertical zones on either side of the center of gravity (47) and a third independent set of two horizontal zones on either side of the center of gravity (47).

24. A system according to claim 23 characterized in that horizontal, vertical, positive and negative slant features in each said segment are computed and output as signature features ((F1..Fn) to the signature database (18) and in that the extracted signature features are each analyzed globally for the complete
10 signature and locally for each segment.

25. A system according to claim 21, 22, 23, or 24 characterized in that two complex element signature features (F121, F122) are extracted by matching and mismatching, a horizontal profile of signature image (26) with a mean horizontal profile of a training sample comprising at least one genuine signature.

26. A system according to claim 25 characterized in that the extracted signature features (52) are
15 evaluated for their distance from a reference obtained from the evaluation program, the distance being computed from at least one genuine signature and being used to effect the authentication decision (17) according to a threshold (63) and in that the threshold (63) is adapted by the evaluation program (14) to enhance the decision.

27. A system according to claim 25 characterized by further comprising an evaluation program (14) to
20 evaluate the results produced by the signature verification program (12) and to provide update information to the signature database (18) for use by signature verification program (12) regarding examined signatures accepted as genuine by signature verification program (12).

28. A system according to claim 27 characterized in that the system comprises a global signature feature database including signature features of genuine and forged samples of signatures of persons other than the
25 specific person's signature together with available genuine samples of the specific person's signatᵤ. ᵣe and in that the global signature database is employed for signature recognition until the database includes a desired number of genuine samples of the specific person's signature.

29. A computer-implemented method of automatic signature authentication comprising:
    a)   examining a digitized signature image (22) obtained from a document, optionally a check,
30   signature image (22) purportedly being the signature of a specific person, while employing an

authenticated signature database (18) of signature features (F1..Fn); and

b)   generating an authentication decision (17) authenticating the examined signature as genuine or
     rejecting the signature as a forgery;

characterized by further comprising:

5   .c)   evaluating the results produced by the signature verification program (12) to provide update
          information to the signature database (18) for use by signature verification program (12) regarding
          examined signatures accepted as genuine by signature verification program (12).

30.   A method according to claim 29 wherein the examined signature is evaluated as genuine and the
database (18) contains less than a desired number of genuine samples characterized by adding the
10   examined sample signature data to the database (18) and recomputing reference statistics for use in
generating the authentication decision (17) using the added examined sample signature data.

31.   A method according to claim 30 characterized in that the reference statistics comprise a signature
feature set and a decision threshold, and by employing a global signature feature database including
signature features of genuine and forged samples of signatures of persons other than the specific person's
15   signature together with available genuine samples of the specific person's signature.

32.   A method according to claim 2 characterized in that the global signature database is employed for
signature recognition until the database includes a desired number of genuine samples of the specific
person's signature.

33.   A method according to claim 31 comprising scanning the digitized signature image (22) from a paper
20   document characterized by generating and examining multiple versions of the signature image (22) and
statistically compiling the results of the examination to enhance the authentication decision (17).

34.   A method according to claim 33 characterized by cleaning digitized signature image (22) prior to
generation of multiple versions to remove printed background characters, noise and unconnected lines not
meeting signature image component criteria.

25   35.   A method according to claim 33 characterized in that the multiple versions of the signature image
(26) comprise an unmodified version (31), a thinned version (32) and a boundary-detected version (40)
and by cleaning signature image (26) to generate the unmodified version (31), passing the unmodified
version (31) directly to a dissimilarity measurement procedure (21), processing a second version of cleaned
signature image (26) in a normalization procedure (33) to adjust the image to a method standardized
30   configuration with regard to orientation and position, thinned version (32) and generating a boundary-
detected version (40) from the normalized second version image.

36. A method according to claim 32, 33, 34 or 35 characterized by defining the signature image (22) into segments, extracting signature features (52) from individual segments and extracting and evaluating signature features (52) against training data received from the evaluation program (14).

37. A method according to claim 9 characterized in that the segments comprise a first set of four
5   quadrants (46) defined from a center of gravity (47) of the signature image, a second independent set of two vertical zones on either side of the center of gravity (47) and a third independent set of two horizontal zones on either side of the center of gravity (47).

38. A method according to claim 36 characterized by extracting and outputting horizontal, vertical, positive and negative slant features in each said segment as signature features ((F1..Fn) to the signature
10   database (18), and analyzing each extracted signature feature globally for the complete signature and locally for each segment.

39. A method according to claim 38 characterized by extracting two complex element signature features (F121, F122) by matching and mismatching, a horizontal profile of signature image (26) with a mean horizontal profile of a training sample comprising at least one genuine signature.

STATEMENT UNDER ARTICLE 19(j)

Collot et al. USP 5,042,073 ("Collot etal.") teaches a computerized signature
verification system comprising a verification algorithm or program to examine a
digitized signature image IS obtained from a document, especially a check, being the
signature image of a specific person, to identify the authorized signatory and to
generate an authentication decision A/R authenticating the signature as genuine or
rejecting the signature as a forgery. Collot et al.'s program employs a
predetermined number of primary signature features (or parameters) regarding the
signatory's signature from a signature feature database (comprising files FR_1,
$Fr_y...FR_{ij}$, module 4D, Fig. 2). Collot et al. also teaches selecting a variable number of
parameters $b1_i$ to bQ_i from the available parameters to minimize decision errors
(column 2, lines 4-9) and storing the selected parameters $b1_i$ to bQ_i in optimum
parameter files (module 4E).

Desired selected parameters $b1_i$ to bQ_y, and certain calculated data derived from the
selected parameters, are supplied by modules 4E and 4F to a decision module 4G
which accepts or rejects the examined signature (column 10, lines 53-58).

The data for primary parameter files (Fp_i), module 4D and optimum reference
parameter files, module 4E are compiled in a learning phase, position PA of switch
4B, module 4C

One problem with Collot et al. is the need for multiple, e.g. 10 (column 4, lines 35-37)
genuine samples of each signature during the learning phase. Frequently, banks and
other users of such signature verification systems may have only one, or possibly
two, such genuine signature samples available.

To solve this problem, the invention of amended system claim 1 and new method
claim 29 provide an evaluation program (14) and method, respectively, to evaluate

the results produced by the signature verification program (12) and to provide update information regarding examined signatures determined to be genuine to the signature database (18) for use by signature verification program (12). By capturing additional genuine signatures samples as they are examined the inventive system is able to increase the number of learning samples in database 18, increasing the accuracy of the system for the signatures having added samples until a desired number of samples, for example 10, is available for a given signature. The claimed system can accordingly be readily utilized by banks and other entities, having only a few, or even a single genuine signature sample.

Collot et al neither teaches nor suggests an evaluation program which feeds back new genuine signature samples to the signature feature database. Thus, amended claim 1 clearly has novelty and inventive step over Collett et al.
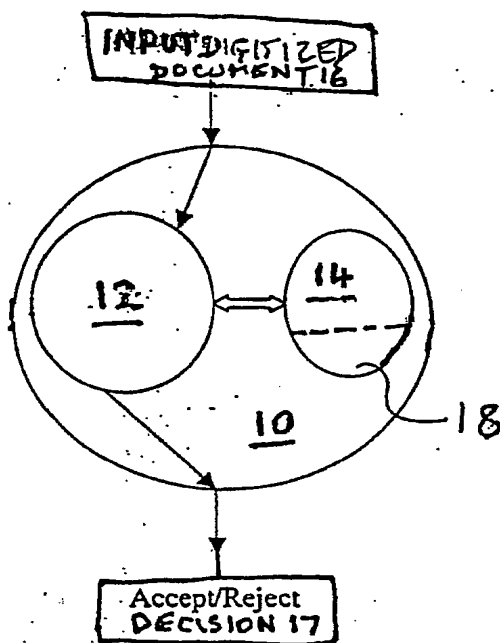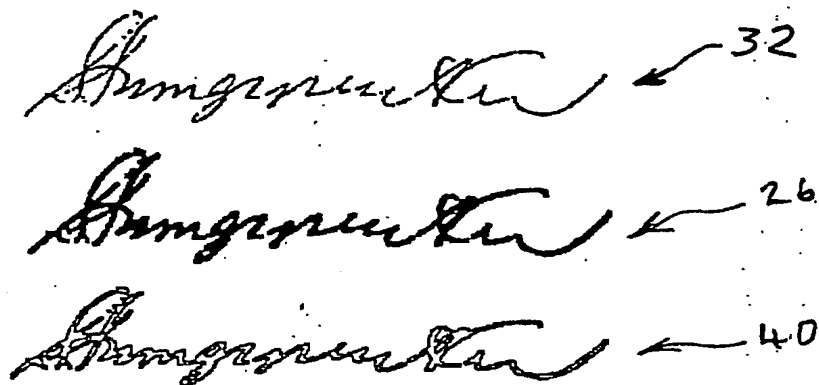
Collot et al's reference data from which decision threshold S is calculated (column 10, 62-65) appears to be separately determined for each signature using only genuine samples of the signature, in a learning phase with switch 4b set to PA, column 4, lines 16 on, especially lines 35-39. No provision is made for updating decision threshold S from additional available signatures. Furthermore, Collot et al. clearly does not appreciate the value of including forgery samples in determining a global threshold value and feature set.

Nor is the usefulness of a global threshold, statistically determined from all available signature samples, in providing reference data for verifying a signature from only one or a few genuine samples.

Claim 20 relates to a system wherein multiple versions of the signature image are generated and examined, which is not remotely suggested by Collot et al.

Döhle et al. employs grey scale value as signature feature recognition elements.

25

Grey scale values impose a significant processing load in view of the need to track data values for a large number of picture elements (column 3, lines 10-28).

Lee et al. requires a digitizer pad (claim 1, lines 65-67) and capture of signature timing information to track dynamic signature features (abstract; claim 1, lines 16-19), neither of which complications is required by the invention of claim 1. Also, Lee et al. requires many genuine signature samples for database creation (column 9, lines 24-25).

Like Collot et al, neither Döhle et al. nor Lee et al. teaches or suggests an evaluation program which feeds back new genuine signature samples to the signature feature database.
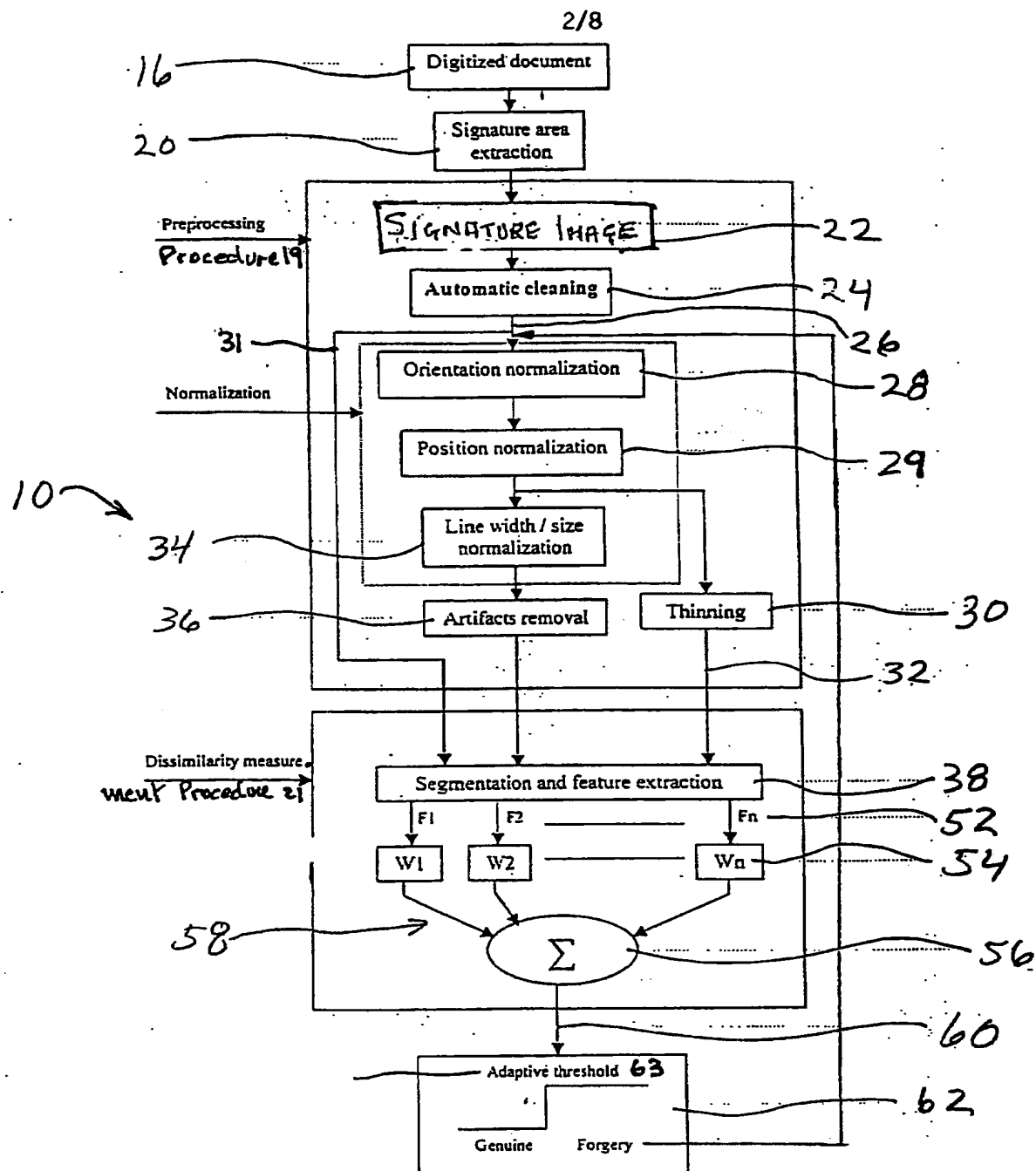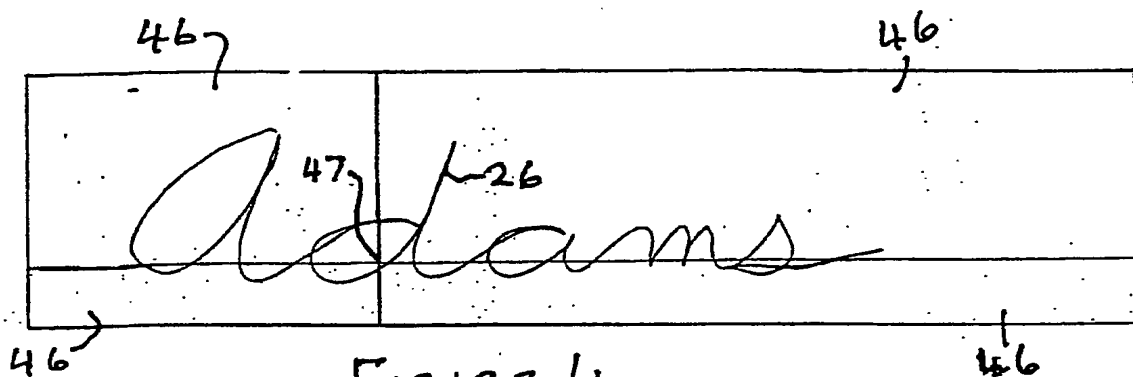
Figure 1

Figure 3

Digitized document

Signature area extraction

Preprocessing
Procedure 19

SIGNATURE IMAGE

Automatic cleaning

Orientation normalization

Normalization

Position normalization

Line width / size normalization

Artifacts removal

Thinning

Dissimilarity measure
ment Procedure 21

Segmentation and feature extraction

| F1 | F2 | | Fn |
| W1 | W2 | | Wn |

Σ

Adaptive threshold **63**

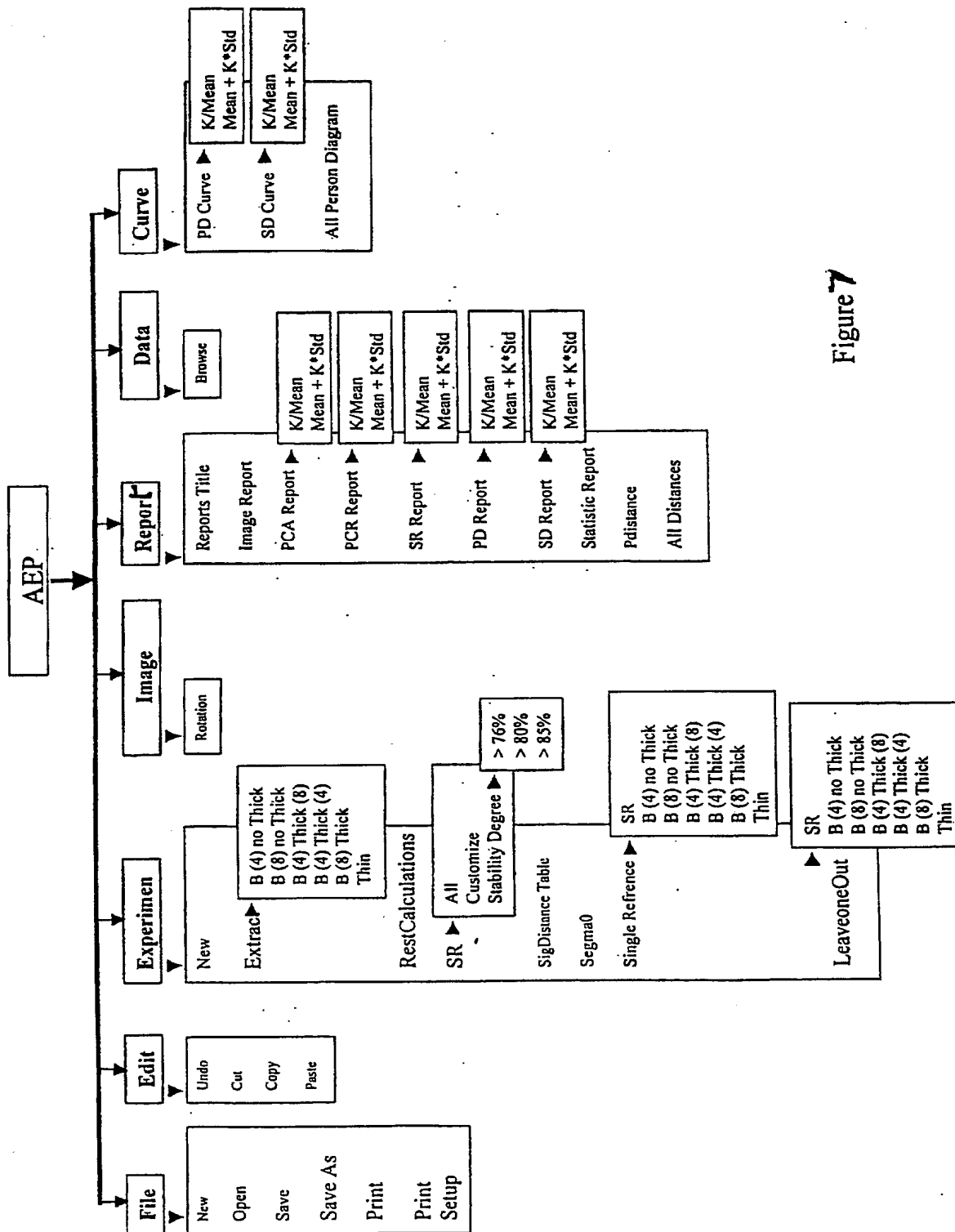Genuine   Forgery
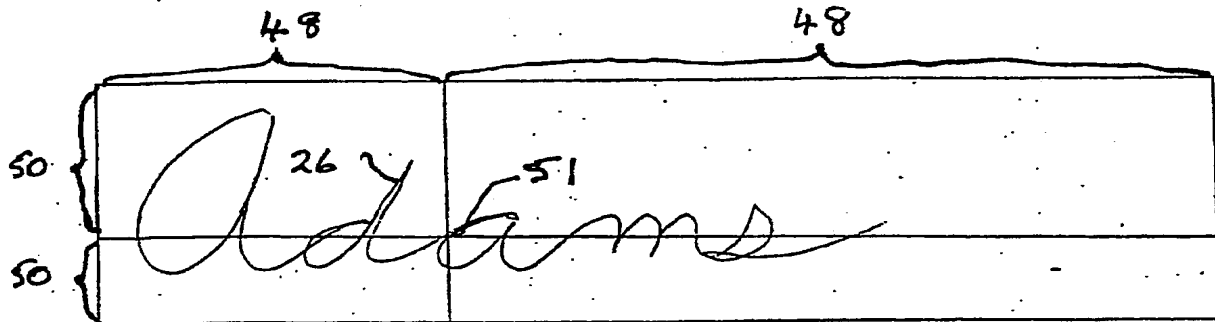
Figure 2
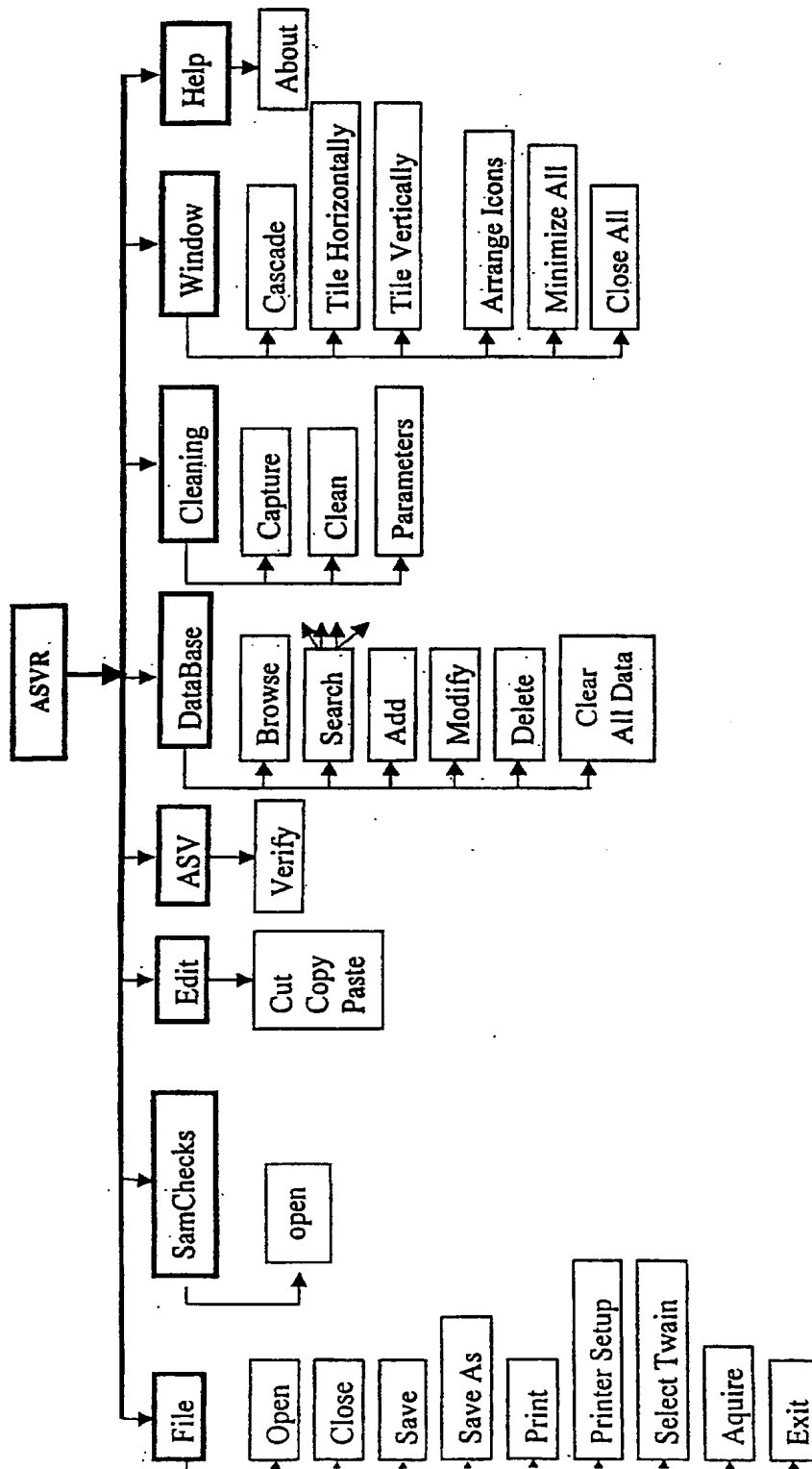
FIGURE 4

FIGURE 5
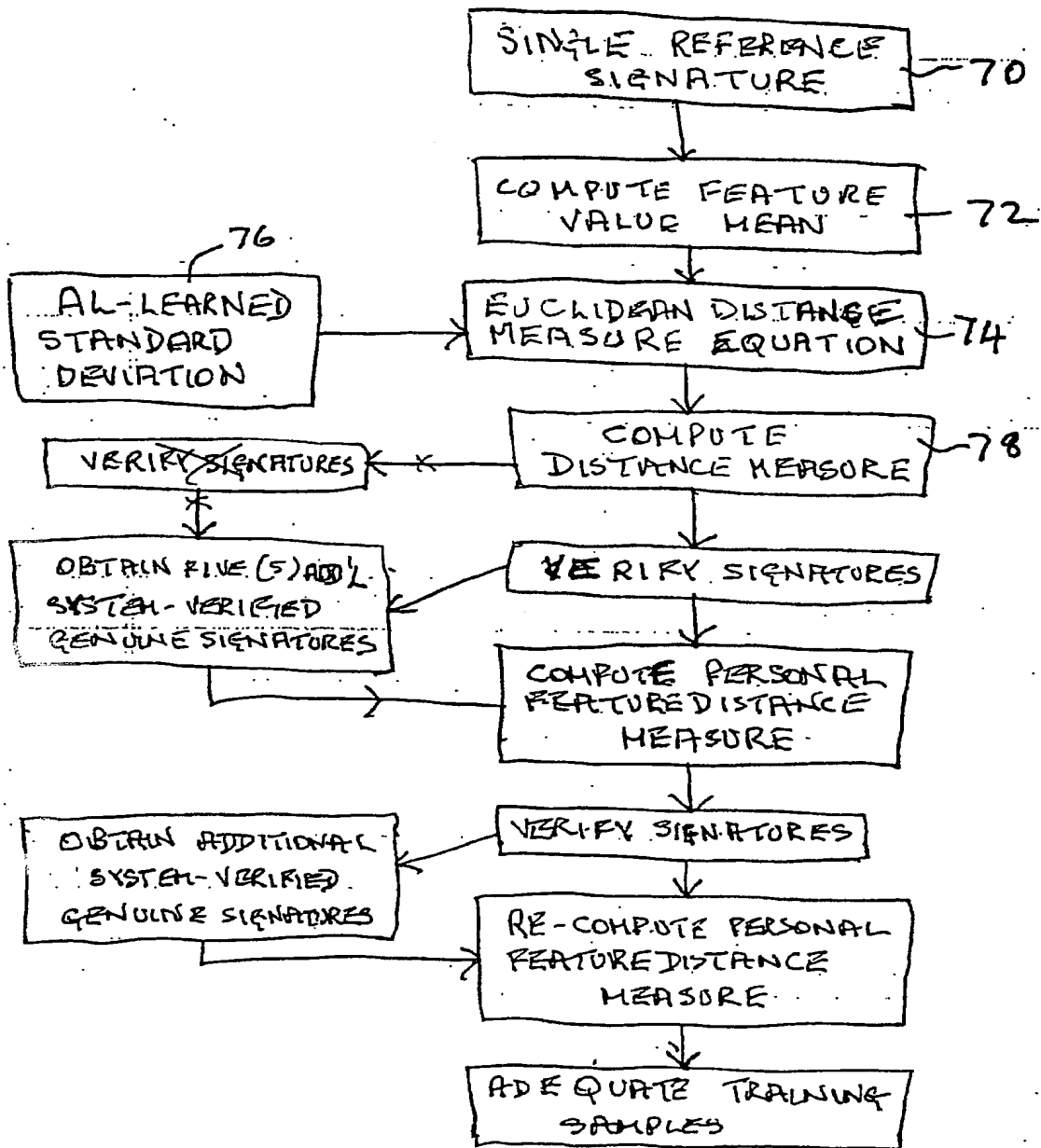
Figure 6

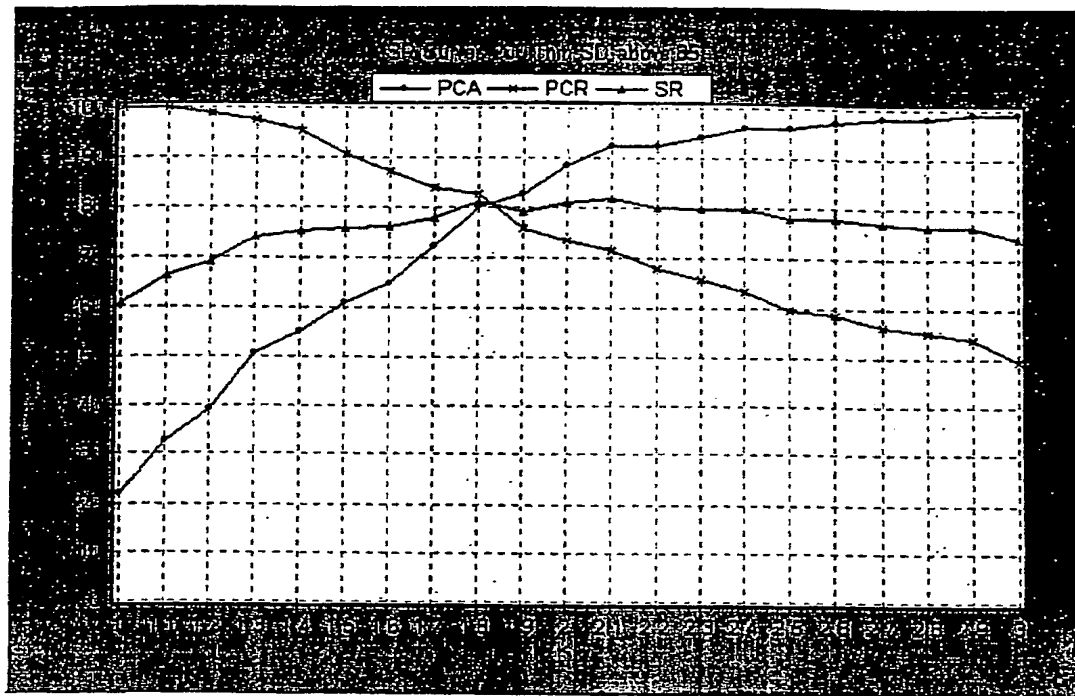Figure 7

ACCELERATED SIGNATURE LEARNING METHOD

FIGURE 8

FIGURE 9

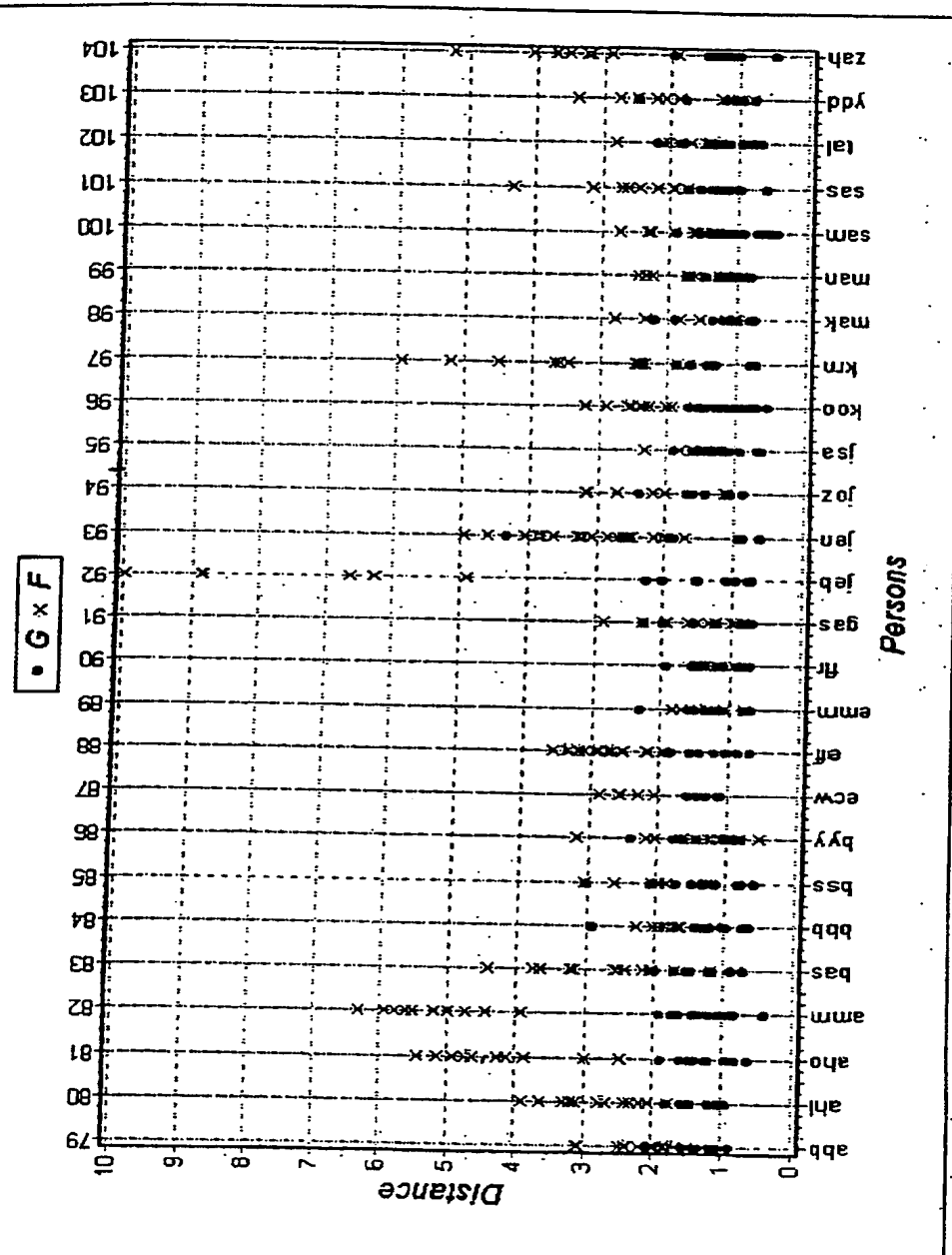Figure 10

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7)   : G06K 9/80

US CL   :382/119

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :   382/116, 119-123; 340/5.52, 5.53, 5.81-5.83; 356/71; 902/3-6; 348/161

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

IEEE ABSTRACTS, US PATENT TEXT FILE USING EAST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,042,073 A (COLLET et al.) 20 August 1991, figures 2 and 16, and column 10, lines 35-68.	11
X	US 5,559,895 A (LEE et al.) 24 September 1996, figure 8, column 12, lines 19-22, column 5, line 60 and column 11, line 21.	1, 9
X	US 5,251,265 A (DOHLE et al.) 05 October 1993, figures 1-3, column 4, lines 42-66 and column 11, lines 10-24.	1-8, 10
A	US 4,028,674 A (CHUANG) 07 June 1977, figures 7 and 13.	3, 10

☐ Further documents are listed in the continuation of Box C.   ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 FEBRUARY 2001	23 FEB 2001

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	BRIAN P. WERNER
Facsimile No. (703) 305-3230	Telephone No. (703) 305-3800

Form PCT/ISA/210 (second sheet) (July 1998) ★